Maximal sub-triangulation as pre-processing phylogenetic data

Anne Berry[*]   Alain Sigayret[*]   Christine Sinoquet[†]

Abstract: *In order to help infer an evolutionary tree (phylogeny) from experimental data, we propose a new method for pre-processing the corresponding dissimilarity matrix, which is related to the property that the distance matrix of a phylogeny (called an* additive matrix*) describes a sandwich family of* chordal *graphs. As experimental data often yield distance values which are known to be under-estimated, we address the issue of correcting the data by* increasing *the distances which are incorrect. This is done by computing, for each graph of the sandwich family, a maximal chordal subgraph.*

1   Introduction

Inferring *evolutionary trees* (also called *phylogenies*) from dissimilarity data remains one of the major challenges in the field of computational biology. Problems like multiple sequence alignment, gene function prediction and protein structure prediction involve phylogenetic reconstruction.

Dissimilarity matrices are obtained experimentally by considering a set of 'objects', which in a phylogenetic context are *taxa*, (but which can, for other biological problems, be genes or proteins for example), and by measuring, using some feasible criterion, the distance between the elements of each pair of taxa. Data are thus described as positive-valued symmetric matrices; when this matrix does indeed correspond to a phylogeny (it is then called an *additive matrix*), reconstructing the tree is easy, and can be done in polynomial time, yielding a unique tree topology ([2], [13], [23]).

Experimental results, however, are not exactly additive matrices, and the phylogeny has to be inferred from real data. Many methods have been proposed for this (see e.g. [9], [13], [24], [26]), but they remain costly and inaccurate, so that new approaches are still being sought.

One of the recent trends in this field of research is to examine the *ordinal properties* of the matrix. Dissimilarity matrices describe a succession of thresholds, and it turns out to be interesting to examine the *structure* of these thresholds, rather than only the values themselves (see [21], [24], [26]), partly because they seem to be less sensitive to small data variations (see [11]), but also because biologists are more concerned with reconstructing the structure (topology) of the phylogenies than by finding the exact valuations of the edges of the phylogeny, so that it seems promising to look for structural properties.

Our approach here is to consider the family of undirected graphs defined by the dissimilarity matrix, each graph of the family corresponding to one of the thresholds of the matrix; we call this the *threshold family of graphs defined by a dissimilarity matrix*.

Huson, Nettles and Warnow in [24] use the property that if the matrix is additive, all the graphs of the threshold family are *chordal* (or *triangulated*), and give experimental evidence that the graphs obtained in real-world data are "almost triangulated".

As a means of pre-processing the experimental data, our aim here is to correct each graph of this threshold family so that it will indeed be chordal, a process called *triangulation*.

The most classical way of correcting a non-chordal graph is called *minimal triangulation*. It is well studied ([27, 3, 4, 7, 6]) and consists in adding an inclusion-minimal set of edges to the graph in order to make it chordal. For a given graph with n vertices and m edges, computing such a minimal triangulation can be done in $O(nm)$ time.

Adding edges to a graph of a threshold family means *lowering* the thresholds of the corresponding edges. Biologists, however, estimate that in the context of constructing a phylogeny from DNA sequences, the thresholds obtained by experimentations tend to be rather too low than too high, the argument being that the number of mutations, represented by the distance between taxa, is in fact higher than what can be observed experimentally (see [26]).

---

[*]LIMOS (Laboratoire d'Informatique de Modélisation et d'optimisation des Systèmes), CNRS UMR 6158, Université Clermont-Ferrand II, Ensemble scientifique des Cézeaux, F-63177 Aubière Cedex, France, berry@isima.fr, sigayret@math.univ-bpclermont.fr.

[†]IRIN (Institut de Recherche en Informatique de Nantes), BP 92208, 44322 Nantes Cedex 3, France, sinoquet@irin.fr.

We have thus examined the problem of correcting a graph which fails to be chordal by *removing* edges rather than adding them, thereby computing a *maximal chordal subgraph* or *maximal subtriangulation* of each of the graphs of the threshold family. The maximal subtriangulation problem has been somewhat less studied than minimal triangulation ([1, 15, 16, 17, 30]), but there exist several algorithms ([1, 16, 30]) which compute a maximal subtriangulation in $O(\Delta m)$ time, where Δ is the maximum degree in the graph.

Our contribution here is a greedy algorithm which will propose a corrected matrix which *increases* the values of any edge on which an anomaly is detected. Our process adds the edges one by one, beginning with an independent set and ending with a clique, in an order as compatible as possible with the input matrix, maintaining throughout a chordal graph at each edge-addition step. This process defines as a side-effect maximal subtriangulations of the graphs defined by the input matrix, but with a better complexity than if the chordal subgraphs were computed separately for each graph of the threshold family.

Our process relies on a new characterization of the edges which can be added to a chordal graph without loosing chordality, which in turn yields a new edge-composition scheme on edges which characterizes chordal graphs.

The paper is organized as follows: we give some previous results and definitions in Section 2; in Section 3, we present and study the threshold family of graphs defined by an additive matrix, and define an edge-addition construction scheme for the class of chordal graphs; Section 4 contains our proposed algorithm; in Section 5, we give some experimental insights then go on to conclude with some open questions and perspectives.

2 Preliminaries

We will first give some definitions and properties which will be useful in the rest of the paper.

2.1 Additive matrices

A **dissimilarity** on a finite set X is a function $\delta\colon X^2 \to \mathrm{IR}^+$ such that $\forall x, y \in X$, $\delta(x, y) = \delta(y, x)$. A dissimilarity is represented by a pairwise comparison symmetric matrix. A **distance** is a dissimilarity such that $\forall x, y \in X$, $\delta(x, y) = 0 \iff x = y$ and $\forall x, y, z \in X$, $\delta(x, y) + \delta(y, z) \geq \delta(x, z)$.

A **phylogeny**, or **evolutionary tree**, is an unrooted binary tree with all edges weighted with positive values. We will denote by $\mathcal{L}$ the set of leaves representing the set of taxa. Figure 1 gives a simple example of such a tree.

For $a, b \in \mathcal{L}$, we will denote by $d(a, b)$ the length of the ab-path from a to b in the phylogeny, which gives the evolutionary distance between a and b. This distance is called an **additive distance** and the associated matrix on $\mathcal{L} \times \mathcal{L}$ is called an **additive matrix**. Note that an additive matrix is a special kind of dissimilarity matrix.
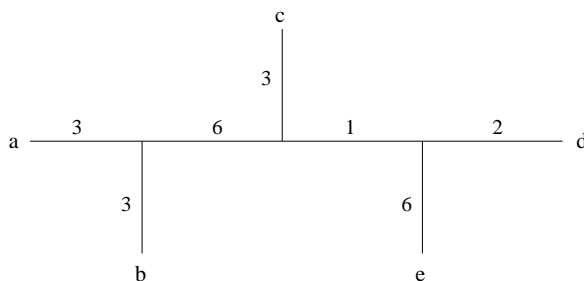

Figure 1: A phylogeny T

Additive matrices are well-studied and the following property, called the Quadrangular Inequality, characterizes them:

Characterization 2.1 *([2]) A distance matrix M on a set of taxa is additive iff for any quadruple $\{a, b, c, d\}$ of taxa, from the 3 sums $d(a, b) + d(c, d)$, $d(a, c) + d(b, d)$ and $d(a, d) + d(b, c)$, the two largest are equal.*

The set of values of a dissimilarity matrix M can be ordered from 0 (as $M[x,x]=0$) to the maximal value. This defines a number of different **thresholds**: 0, 1, ..., k, in increasing order. An **ordinal matrix** of a dissimilarity matrix is thus defined as the matrix obtained by replacing each dissimilarity value by its threshold. We will denote by θ the function giving the threshold rank corresponding to a dissimilarity.

Example 2.2 *The phylogeny from Figure 1 yields the following dissimilarity and ordinal matrices:*

M	a	b	c	d	e
a	0	6	12	12	16
b		0	12	12	16
c			0	6	10
d				0	8

The dissimilarity matrix M of T

W	a	b	c	d	e
a	0	1	4	4	5
b		0	4	4	5
c			0	1	3
d				0	2

The ordinal matrix W of T

Note that M is an additive matrix.

The corresponding 6 thresholds are the following: $\theta(0)=0$, $\theta(6)=1$, $\theta(8)=2$, $\theta(10)=3$, $\theta(12)=4$, $\theta(16)=5$. Dissimilarity values are: $\theta^{-1}(0)=0$, $\theta^{-1}(1)=6$, $\theta^{-1}(2)=8$, $\theta^{-1}(3)=10$, $\theta^{-1}(4)=12$, $\theta^{-1}(5)=16$.

2.2 Chordal graphs and triangulations

A graph $G=(V,E)$ is said to be **chordal** or **triangulated** if it contains no chordless cycle on more than 3 vertices. We will need the following tree-oriented characterization for chordal graphs, due independently to Walter, Buneman and Gavril:

Characterization 2.3 *([29], [14], [19]) A graph is chordal iff it is the intersection graph of a family of subtrees of a tree.*

Graph inclusion: If $G=(V,E)$ is a graph and $G'=(V,E')$ is another graph on the same vertex set, we will write $G \subseteq G'$ iff $E \subseteq E'$ and $G \subset G'$ iff $E \subset E'$ ($\subset$ denotes strict inclusion).

In their famous paper, Rose Tarjan and Lueker gave the following definition of minimal triangulation:

Definition 2.4 *([27]) If $G=(V,E)$ is a non-chordal graph, a chordal graph $H=(V,E+F)$ is said to be a minimal triangulation of G if $\forall F' \subset F$, graph $(V,E+F')$ fails to be chordal.*

In the same paper, they proved the fundamental result that only one edge needs to be removed and the resulting graph tested:

Theorem 2.5 *([27]) Let $G=(V,E)$ be a graph, let $H=(V,E+F)$ be a chordal graph; H is a minimal triangulation of G iff $\forall f \in F$, graph $(V,(E+(F \setminus \{f\})))$ fails to be chordal.*

This result relies on the following very important Lemma from the same paper, which ensures that, given two chordal graphs which are mutually inclusive, there is an ordering on the edges which need to be added to the smaller graph which will maintain chordality at each edge-addition step.

Lemma 2.6 *([27]) Let $G_1=(V,E_1)$ be a chordal graph, let $G_2=(V,E_2)$ be a chordal graph such that $G_1 \subset G_2$. Then $\exists f \in E_2 \setminus E_1$ such that $G'=(V,E_2 \setminus \{f\})$ is chordal.*

We would like to point out that the property described by Lemma 2.6 is far from trivial; it fails to hold for hole-free graphs (graphs with no chordless cycle with length strictly greater than 4), as illustrated by the counterexample below, and it is not known whether it holds for weakly chordal graphs (graphs with no hole, and no hole in the complement).

Counterexample 2.7 *Graphs G_1 and G_2 in figure 2 are hole-free graphs. G_2 can be obtained from G_1 by adding edges ac and df, but $G_1+\{ac\}$ and $G_1+\{df\}$ are not hole-free graphs.*

Maximal subtriangulation was to the best of our knowledge introduced in 1983 by Erdös and Laskar ([17]) in view of removing a minimum number of edges in order to make a graph chordal. Maximal subtriangulation is defined in a fashion similar to minimal triangulation:

Definition 2.8 *Let $G=(V,E)$ be a non-chordal graph, let $H=(V,E \setminus F)$ be a chordal graph. We will say that H is a maximal sub-triangulation of G if $\forall F' \subset F,(V,(E \setminus F)+F')$ fails to be chordal.*
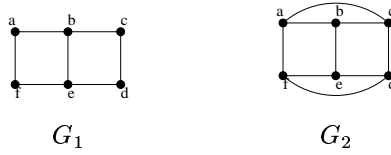
Figure 2: Graphs of Example 2.7

3 Maintaining a family of chordal graphs

3.1 The threshold family of graphs defined by a dissimilarity matrix

We will use the ordinal matrix associated with a dissimilarity matrix to define the corresponding threshold family of graphs:

Definition 3.1 *Let $\mathcal{A}$ be a set of taxa with dissimilarity matrix M; let W be the corresponding ordinal matrix, on thresholds 0, 1, ..., k. We will define a family of graphs $G_0 \subset G_1 \subset ... \subset G_k$, called the* **threshold family of graphs** *associated with W (and thus with M), with $G_i = (V, E_i), V = \mathcal{A}$, and $ab \in E_i$ iff $W_{\mathcal{A}}[a, b] \leq$ i.*

Remark 3.2 The threshold family of graphs defined above should not be confused with threshold graphs, which are defined in correlation with integer programming; threshold graphs are characterized as being chordal, with a chordal complement, and P_4-free (see[20] and [12]), which is definitely not the general case for the graphs of a threshold family as defined by Definition 3.1.

Note that G_0 is an independent set (a graph with no edges) and that G_k is a clique (a graph with all possible edges).

The threshold matrix W induces on the set of edges $V \times V$ a preorder relation $\mathcal{R}$: $ab\,\mathcal{R}\,cd$ iff $W[a, b] \leq W[c, d]$. $\mathcal{R}$ defines an *ordered partition of the edges* of G_k; each class F_i of edges is defined by $F_i = E_i - E_{i-1} = \{xy \mid W[x, y] =$ i $\}$. Graph G_i is obtained from graph G_{i-1} by adding set of edges F_i. $\mathcal{R}$ defines a total ordering on theses classes, with $F_i < F_j$ iff i $<$ j.

Example 3.3 *Figure 3 illustrates the family of non-trivial graphs constructed from the matrix of Example 2.2.*
Ordered partition on the edges: $\{ab, cd\} < \{de\} < \{ce\} < \{ac, ad, bc, bd\} < \{ae, be\}$
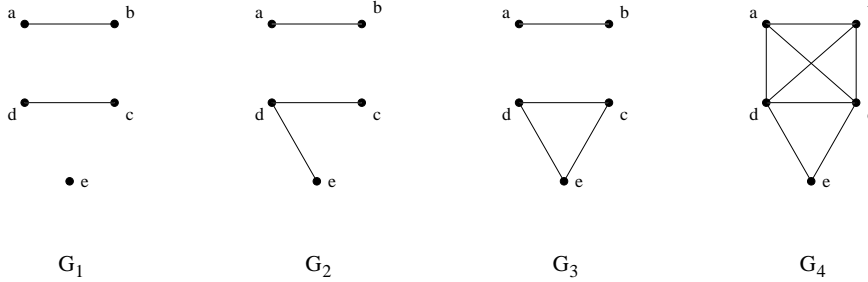

Figure 3: Graphs $G_1 \subset G_2 \subset G_3 \subset G_4$ representing matrix of Example 2.2. G_0 is an independent set and G_5 is a clique.

Property 3.4 *If M is an additive matrix then the threshold family of graphs defined by M is a family of chordal graphs.*

Proof: Let T be the phylogeny associated with an additive matrix M, let G_i be the graph corresponding to threshold i $\in [0..k]$. It is easy to add extra internal nodes to T in order to obtain a tree T' where there is a node at mid-distance between any pair $\{a, b\}$ of vertices. Let us now consider the family of subtrees of T' defined by: for each leaf x, T'_x is the subtree containing all nodes at distance $\theta^{-1}($i$)/2$ or less from x; G_i is the interrsection graph of this family. By virtue of Characterization 2.3, G_i is chordal. ◇

4

The converse of Property 3.4 fails to be true: there are ordinal matrices which represent a chordal threshold family of graphs, but for which there is no corresponding additive matrix.

Counterexample 3.5 *An ordinal matrix W which can be associated with no dissimilarity matrix:*

W	a	b	c	d
a	0	3	2	1
b		0	4	4
c			0	1

W *is associated with a threshold family of chordal graphs, but this matrix is the ordinal matrix of no additive matrix: for any dissimilarity δ of which W is the ordinal matrix, we have $\delta(a,c) + \delta(b,d) = \theta^{-1}(2) + \theta^{-1}(4)$, but $\delta(a,b) + \delta(c,d) = \theta^{-1}(3) + \theta^{-1}(1)$ and $\delta(a,d) + \delta(b,c) = \theta^{-1}(1) + \theta^{-1}(4)$, so $\delta(a,c) + \delta(b,d)$ is strictly greater then the other two sums, which contradicts the Quadrangular Inequality 2.1.*

3.2 Preconditionning matrices

Experimental results show that not only do the dissimilarity matrices biologists have to work with fail to be additive, but the corresponding graphs very often fail to be chordal.

As we have stated in our introduction, our goal is to precondition a matrix into describing a family of chordal graphs, while dealing with threshold values which are too low. As a result of Counterexample 3.5, forcing a threshold family into a chordal family will not in general be sufficient to ensure that the matrix becomes additive; however, it may well be an important first step in error recovery.

Example 3.6 *Modified dissimilarity matrix of Example 2.2, with "incorrect" values for ad and bc, which have been lowered from 12 to 8. This describes a family (H_i) of graphs. H_2 and H_3 fail to be chordal.*

	a	b	c	d	e
a	0	6	12	8	16
b	/	0	8	12	16
c	/	/	0	6	10
d	/	/	/	0	8
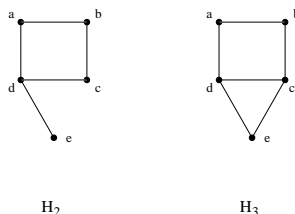
3.3 An edge-addition composition scheme for chordal graphs

In order to compute a threshold family of graphs which are chordal and such that each graph G'_i of the new family is a subgraph of the corresponding original graph G_i, we will construct clique G_k from independent set G_0 by adding at each step an inclusion-maximal set of edges which maintains chordality.

The problem of maintaining a chordal graph while adding edges has been examined by Ibarra ([25]), who has obtained good results on the queries as to whether inserting or deleting an edge preserves a chordal graph. He uses clique trees (see [10] for a solid introduction of this concept) as an intermediate representation, and uses his work as an illustration of the power of this representation of a chordal graph. He derives a characterization of the edges which can be inserted or deleted, expressed in terms of clique trees.

We will propose yet a different approach, for which we will need the notion of 2-pair, defined by Hayward, Hoàng and Maffray to characterize weakly chordal graphs.

Definition 3.7 *([22]) A pair $\{a, b\}$ of non-adjacent vertices is called a* 2-pair *iff every chordless path from a to b is of length exactly 2.*

Theorem 3.8 *Let G_1 be a chordal graph, let $\{a, b\}$ be a pair of non-adjacent vertices of G_1, let G_2 be the graph obtained from G_1 by adding edge ab; then G_2 is chordal iff $\{a, b\}$ is a 2-pair of G_1.*

Proof: Let G_1 be a chordal graph, let $\{a, b\}$ be a pair of non-adjacent vertices of G_1, let G_2 be the graph obtained from G_1 by adding edge ab, let $\mu = ax_1x_2...x_kb$ be a longest chordless path from a to b in G_1. In G_2, $ax_1x_2...x_kba$ will be a chordless path on more than 3 vertices iff μ is of length greater than 2, that is iff $\{a, b\}$ fails to be a 2-pair of G_1. $\diamond$

An additional property which is vital to our problem is ensuring that we are able to move from one graph of the threshold family to the next.

Property 3.9 *Let G_1 be a chordal graph, let G_2 be a chordal graph such that $G_1 \subset G_2$. Then G_2 can be obtained from G_1 by repeatedly adding an edge between the two vertices forming a 2-pair.*

Proof:
Let G_1 be a chordal graph, let G_2 be a chordal graph such that $G_1 \subset G_2$. By Lemma 2.6, $\exists ab \in E_2 \setminus E_1$ such that $(V, E_2 \setminus \{ab\})$ is chordal. By Theorem 3.8, $\{a,b\}$ is a 2-pair of $G_2 \setminus \{ab\}$. If we repeat this until we obtain graph G_1, we will have constructed (in reverse) a 2-pair edge addition ordering which enables us to construct G_2 from G_1. ⋄

We use Theorem 3.8 to propose the following composition scheme for chordal graphs, which starts with an independent set and constructs the desired chordal graph by an edge-addition process. This is well-adapted to our problem, as we want to start with an independent set, scan a succession of mutually inclusive chordal graphs, and end with a clique, which is also chordal.

Composition Scheme 3.10 *A graph on n vertices is chordal iff it can be constructed by starting with an independent set on n vertices, and by adding at each step an edge between the two vertices forming a 2-pair.*

Remark 3.11 *Classical composition schemes for chordal graphs are vertex-addition schemes, such as starting with a clique and adding a simplicial vertex at each step. Very recent work by Berry, Heggernes and Villander ([6]) gives a much more general process for adding a vertex v to a chordal graph, by characterizing the edges incident to v which must be added along with an edge vw in order to maintain chordality.*

4   An additive data pre-processing algorithm

4.1   Algorithmic strategy

We now propose an algorithm based on Composition Scheme 3.10, which uses as input a dissimilarity matrix M and outputs a dissimilarity matrix M' defining a threshold family of chordal graphs, and which raises all the thresholds it modifies.

Our algorithm starts with an independent set of vertices (graph G_0), and at each step i will construct graph G_i from graph G_{i-1} by adding as many edges as possible. The algorithm at step i repeatedly chooses, from a set of candidate pairs, a pair of vertices which allows to maintain a chordal graph.

At the beginning of step i, a candidate pair is defined as any pair $\{a,b\}$ such that $M[a,b] \leq \theta^{-1}(i)$ and ab is not an edge of G_{i-1}.

In order to remain as close as possible to the original matrix, we will give priority to the candidate pairs which correspond to the smallest threshold. We will implement this by using a FIFO queue; at each step, i, the new candidate pairs are added to the queue, and the algorithm then repeatedly chooses the first pair of the queue which is a 2-pair of the current graph, and adds it to the current graph G_i under construction.

By Property 3.9, at the end of the algorithm, the FIFO queue is empty and every edge has been given a threshold in the corrected matrix M' obtained.

4.2   Algorithm

Algorithm ADD-SUB-TRI
Input: A dissimilarity matrix M on n taxa, with threshold 0, ..., k.
Output: An dissimilarity matrix M', such that every graph in the threshold family is chordal.
Initialization:
G_0 is an independent set on n vertices;
Create an empty FIFO queue Q;
begin
For i=1 **to** k−1 **do**
    $G_i \leftarrow G_{i-1}$;
    Compute the set F_i of pairs $\{a,b\}$ such that $M[a,b] = \theta^{-1}(\text{i})$;

Add F_i to Q;
Repeat
    Scan Q and remove the first pair ab which is a 2-pair;
    Add edge ab to graph G_i;
    $M[a, b] \leftarrow \theta^{-1}(\texttt{i})$;
    Until Q contains no 2-pair of G_i;
Give all remaining edges in Q value $\theta^{-1}(\texttt{k})$ in M';
Add all remaining edges in Q to G_{k-1} to form G_k, a clique on n vertices.
end

Example 4.1 *On the "incorrect" matrix given in Example 3.6, at step 2, adding edge bc after adding edge ad would induce a 4-cycle abcd . We will add edge bc at step 4, after edge bd, thus raising the value of bc from 8 back to its "normal" value, 12. Note that edge ad has not been corrected.*

As a consequence of Composition Scheme 3.10, Algorithm ADD-SUB-TRI computes a threshold family of graphs of which each member is a maximal sub-triangulation of the corresponding graph of the original matrix.

4.3   Complexity Analysis

In [28], Spinrad and Sritharan propose an algorithm which repeatedly adds a 2-pair to the graph; they use a data structure which maintains the "2-pair structure" of the graph, which costs $O(n)$ to update for each edge addition. As there are $O(n)$ edges to process, using this 2-pair structure, our global complexity is thus $O(n^4)$.

Note that if we computed a maximal subtriangulation in $O(\Delta m)$ time for each of the $O(n^2)$ graphs of the threshold family, this would cost $O(n^5)$.

5   A few experimental results

We have implemented Algorithm ADD-SUB-TRI and run some preliminary tests.

We first ran it on real data to measure how "distant" a matrix obtained experimentally can be from a triangulated distance matrix (a matrix associated to a family of chordal graphs). A set of 5 matrices, ranging from size 11 to 57 (quite current sizes for phylogeny reconstruction) and dealing with plants and bacteria, presents a percentage of triangulated graphs varying widely, from 4% to 96%. We ran ADD-SUB-TRI on these, and used existing reconstructing algorithms on both the initial matrix and the output matrix; we found no significant difference between the two results, only a slightly better concensus in two cases.

We then started testing ADD-SUB-TRI on artificial data.

Our experimental protocol is the following : we begin by randomly generating an additive matrix A, then by randomly generating from A a biased matrix B obtained by decreasing some of the dissimilarity values and by varying the bias; finally, we run Algorithm ADD-SUB-TRI on B, resulting in matrix C, and we use various metrical and topological criteria, such as those described in [18], to evaluate whether C is nearer to A then B is.

This was run on a hundred matrices of size 20. The results, described in [8], show that this algorithm is indeed promising as a pre-processing technique for improving phylogenetic data: in roughly one third of the cases encountered in our simulations, the reconstruction was better using improved matrix C than input matrix B using standard phylogenetic reconstruction algorithm; in all remaining cases, the reconstruction was the same.

Because of these promising preliminary results, we aim to pursue experimentation on a larger scale with artificial data, and also to examine more real data.

6   Conclusion, perspectives and open questions

Regarding the complexity of Algorithm ADD-SUB-TRI as presented in this paper, we use data structures from [28], which deals with the problem of maintaining a 2-pair structure in an arbitrary graph, whereas we deal with chordal graphs only. We believe that for chordal graphs this complexity should be improved, especially so since in a chordal graph there are many 2-pairs which are not disrupted by an edge-addition, so that it may not be necessary to update the 2-pair structure at every edge-addition step.

We also feel that in many biological data preprocessing problems, such as those dealing with biochip data, it may be interesting to maintain a chordal graph, but not necessarily by systematically lowering existing threshold; it may be interesting to use the process described in [6] to allow the user to choose at each step whether to lower or to raise the thresholds, depending on how many modifications this causes.

References

[1] E. Balas. *A fast algorithm for finding an edge-maximal subgraph with a TR-formative coloring.* Discrete Applied Mathematics, 15, 123–134, 1986.

[2] J-P. Barthélémy and A. Guénoche. *Trees and Proximity Representations.* J. Wiley and Sons (eds.), 1991.

[3] A. Berry. *A wide-range efficient algorithm for minimal triangulation.* Proc. tenth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'99), 860–861, 1999.

[4] A. Berry, J. Blair, and P. Heggernes. *Maximum cardinality search for computing minimal triangulations.* In Graph Theoretical Concepts in Computer Science - WG 2002, LNCS 2573, L. Kucera (ed.), Springer Verlag, 1–12, 2002.

[5] A. Berry, J.-P. Bordat, and P. Heggernes. *Recognizing weakly triangulated graphs by edge separability.* Nordic Journal of Computing, 7, 164–177, 2000.

[6] A. Berry, P. Heggernes, and Y. Villanger. *An on-line incremental approach for dynamically maintaining chordal graphs.* Research Report LIMOS: RR 2003-04, 2003.

[7] A. Berry, J-P. Bordat, P. Heggernes, G. Simonet, and Y. Villanger. *A wide-range algorithm for minimal triangulation from an arbitrary ordering.* Technical Report Reports in Informatics 243, University of Bergen (Norway); Research Report LIMOS: RR 2003-02, 2003. *Submitted to Journal of Algorithms.*

[8] A. Berry, A. Sigayret and C. Sinoquet. *Towards improving phylogeny reconstruction with combinatorial-based constraints on an underlying family of graphs.* Research Report LIMOS: RR 02-103, 2002.

[9] V. Berry and O.Gascuel. *Inferring evolutionary trees with strong combinatorial evidence.* Theoretical Computer Science, 240, 2, 271–298, 2000.

[10] J. R. S. Blair, B. Peyton. *An introduction to chordal graphs and clique trees. Graph Theory and Sparse Matrix Computation, 56, 1–29, 1993.*

[11] *F. Bonnot, A. Guénoche, X. Perrier. Properties of an order distance associated with a tree distance. Proceedings of OSDA'95 (Ordinal and Symbolic Data Analysis), E.Diday et al. (eds.), Springer Verlag, 252–261, 1996.*

[12] A. Brandstädt, V. B. Le and J. Spinrad. *Graph classes - a survey.* SIAM Monographs on Discrete Mathematics and Applications, 1999.

[13] P. Buneman. *The recovery of trees from measures of dissimilarity.* Mathematics in the Archeological and Historical Sciences, Edinburgh University Press, 387–395, 1971.

[14] P. Buneman. *A characterization of rigid circuit graphs.* Discrete Mathematics, 9, 205–212, 1974.

[15] T. F. Coleman. *A chordal preconditioner for large-scale optimization.* Applied Mathematics, 40, 265–287, 1988.

[16] P. M. Dearing, D. R. Shier, and D. D. Warner. *Maximal chordal subgraphs.* Discrete Applied Mathematics, 20, 181–190, 1988.

[17] P. Erdös and R. Laskar. *On maximum chordal subgraph.* Cong. Numerantium, 39, 367–373, 1983.

[18] H. Garetta, A. Guénoche. *How confident can we be that a tree representation is good? (Quelle confiance accorder à une représentation arborée?).* Proc. JOBIM 2000, LNCS, O.Gascuel and M.-F. Sagot Eds., Springer Verlag, 2066, 45–56, 2001.

[19] F. Gàvril. *The intersection graphs of subtrees of trees are exactly the chordal graphs.* Journal Combinatorial Theory B, 16, 47–56, 1974.

[20] M. C. Golumbic. *Algorithmic Graph Theory and Perfect Graphs.* Acad. Press, New York, 1980.

[21] A. Guénoche. *Ordinal Properties of tree distances.* Discrete Applied Mathematics, 192, 103–117, 1998.

[22] R. Hayward, C. Hoàng, and F. Maffray. *Optimizing weakly triangulated graphs.* Graphs and Combinatorics, 5, 339–349, 1989.

[23] J. Hein. *An optimal algorithm to reconstruct trees from additive distance data.* Bulletin of Mathematical Biology, 51(5), 597–603, 1989.

[24] D. Huson, S. Nettles, T. Warnow. *Obtaining highly accurate topology estimates of evolutionary trees from very short sequences.* Proc. RECOMB'99, Lyon (France), 198–207, 1999.

[25] L. Ibarra. *Fully dynamic algorithms for chordal graphs and split graphs.* University of Victoria, Technical report, DCS-262-IR, 2000.

[26] P. Kearney, R. Hayward, H. Meijer. *Inferring evolutionary trees from ordinal data.* Proc. eighth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'97), 418–426, 1997.

[27] D. Rose, R.E. Tarjan, and G. Lueker. *Algorithmic aspects of vertex elimination on graphs.* SIAM Journ. Comput., 5, 146–160, 1976.

[28] J. Spinrad and R. Sritharan. *Algorithms for weakly triangulated graphs.* Discrete Applied Mathematics, 59, 181–191, 1995.

[29] J.R. Walter *Representations of Rigid Circuit Graphs.* PhD. Dissertation, Wayne State University, Detroit, 1972.

[30] J. Xue. *Edge-maximal triangulated subgraphs and heuristics for the maximum clique problem.* Networks, 24, 109–120, 1994.